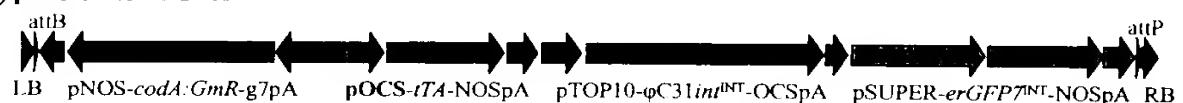


FIGURE 1

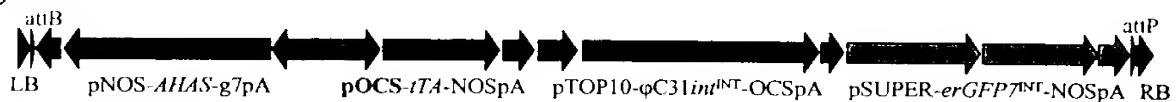
A) pBPS EW051 T-DNA



B) pBPS EW151 T-DNA



C) Monocot T-DNA



D) T-DNA Foot Print



FIGURE 2

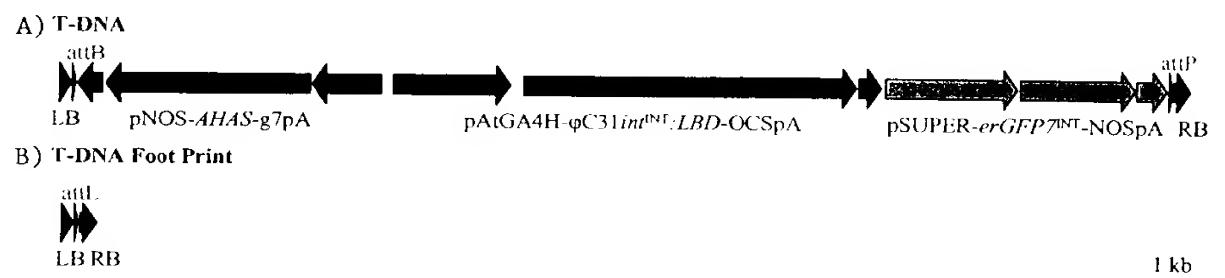
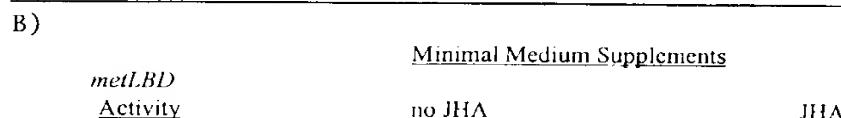
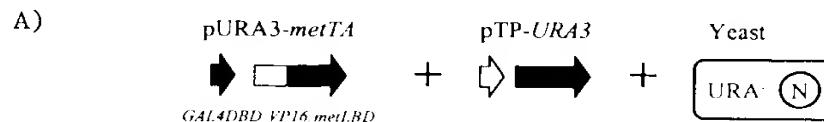


FIGURE 3

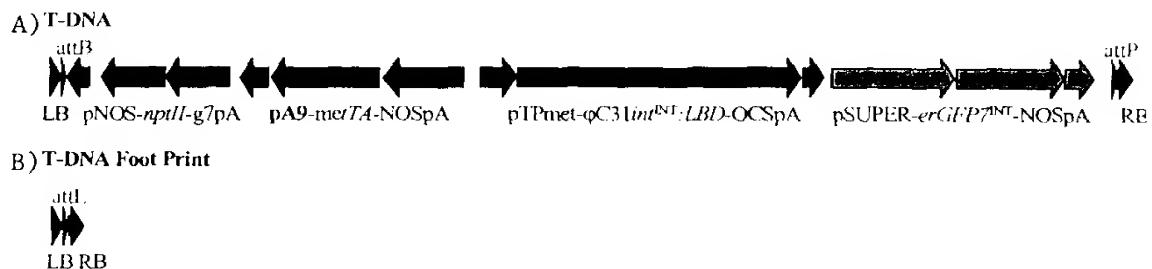


○ = metTA Localization

FIGURE 4



FIGURE 5



## FIGURE 6

### Nucleotide sequence of $\phi$ C31int<sup>INT</sup>

1 ATGGCACAAG GGGTTGTGAC CGGGTGGAT ACGTAAGTTT CTGCTTCTAC CTTTGATATA  
61 TATATAATAA TTATCATTAA TTAGTAGTAA TATAATAATT CAAATATT TTTCAAAATA  
121 AAAGAATGTA GTATATAGCA ATTGCTTTG TGATGTTAT AAGTGTGTAT ATTTAATT  
181 ATAACTTTC TAATATATGA CCAAAATTG TTGATGTGCA GGTAACGGGG TGCTTACGAC  
241 CGTCAGTCGC GCGAGCGCGA GAATTGAGC GCAGCAAGCC CAGCGACACA GCGTAGCGCC  
301 AACGAAGACA AGGCAGGCCA CCTTCAGCGC GAAGTGGAGC GCGACGGGG CCGGTTCAAGG  
361 TTCGTCGGGC ATTCAGCGA AGCGCCGGC ACGTGGCGT TCGGGACGGC GGAGCGCCCG  
421 GAGTCGAAC GCATCCTGAA CGAATGCCGC GCCGGGGGG TCAAACATGAT CATTGTCTAT  
481 GACGTGTCGC GCTTCTCGCG CCTGAAGGTC ATGGACCGA TTCCGATTGT CTCGAAATTG  
541 CTCGCCCTGG GCGTGACGAT TGTTCCACT CAGGAAGGCG TCTTCCGGCA GGGAAACGTC  
601 ATGGACCTGA TTCACCTGAT TATGCGGCTC GACCGCTCGC ACAAAAGAAC TTCGCTGAAG  
661 TCGGCGAAGA TTCTCGACAC GAAGAACCTT CAGCGCGAAT TGGGCGGGTA CGTCGGCGGG  
721 AAGGGCCCTT ACGGCTTCGA GCTTGTTCG GAGACGAAGG AGATCACCGC CAACGGCCGA  
781 ATGGTCAATG TCGTCATCAA CAAGCTTGC CGACTCGACCA CTCCCCCTTAC CGGACCCCTTC  
841 GAGTCGAGC CCGACGTAAT CCGGTGGTGG TGGCGTGAGA TCAAGACGCA CAAACACCTT  
901 CCCTTCAAGC CGGGCAGTC AGCCGCCATT CACCCGGCA GCATCACGGG GCTTTGTAAG  
961 CGCATGGACG CTGACGCCGT GCCGACCCGG GGCAGACGA TTGGGAAGAA GACCGCTTCA  
1021 AGCGCCTGGG ACCCGGCAAC CGTTATGCGA ATCCTTGGG ACCCGCGTAT TGCGGGCTTC  
1081 GCCGCTGAGG TGATCTACAA GAAGAACCG GACGGCACGC CGACCACGAA GATTGAGGGT  
1141 TACCGCATTG AGCGCAGCC GATCACGCTC CGGCCGGTCG AGCTTGTATT CGGACCGATC  
1201 ATCGAGCCCG CTGAGTGGTA TGAGCTTCAG GCGTGGTTGG ACGGCAGGGG GCGGGCAAG  
1261 GGGCTTCCC GGGGGCAAGC CATTCTGTCC GCCATGGACA AGCTGTACTG CGAGTGTGGC  
1321 GCCGTATGA CTTCGAAGCG CGGGGAAGAA TCGATCAAGG ACTCTTACCG CTGCCGTGCG  
1381 CGGAAGGTGG TCGACCCGTC CGCACCTGGG CAGCACGAAG GCACGTGCAA CGTCAGCATG  
1441 CGGGCACTCG ACAAGTTGT TGCGGAACGC ATCTTCAACA AGATCAGGCA CGCCGAAGGC  
1501 GACGAAGAGA CGTTGGCGCT TCTGTGGAA GCGCCCGAC GCTTCGGCAA GCTCACTGAG  
1561 GCGCCTGAGA AGAGCGGCGA ACGGCGAAC CTTGTTGCGG AGCGCGCCGA CGCCCTGAAC  
1621 GCCCTTGAAG AGCTGTACGA AGACCGCGCG GCAGGGCGGT ACGACGGACC CGTTGGCAGG  
1681 AAGCACTTCC GGAAGCAACA GGCAGCGCTG ACGCTCCGGC AGCAAGGGG GGAAGAGCGG  
1741 CTTGCCGAAC TTGAAGCCGC CGAACCCCCG AAGCTTCCCC TTGACCAATG GTTCCCCGAA  
1801 GACGCCGACG CTGACCCGAC CGGCCCTAAG TCGTGGTGGG GCGCGCGTC AGTAGACGAC  
1861 AAGCGCGTGT TCGTCGGGCT CTTCTGAGAC AAGATCGTTG TCACGAAGTC GACTACGGGC  
1921 AGGGGGCAGG GAACGCCAT CGAGAACGCGC GCTTCGATCA CGTGGGCGAA GCCGCCGACC  
1981 GACGACGACG AAGACGACGC CCAGGACGGC ACGGAAGACG TAGCGGGCGTA G

## FIGURE 7

### Nucleotide sequence of $\varphi C31 int^{*int}$

1 ATGGCACAAG GGGTTGTGAC CGGGGTGGAT ACGTAAGTTT CTGCTTCTAC CTTTGATATA  
61 TATATAATAA TTATCATTAA TTAGTAGTAA TATAATATT CAAATA TTTT TTTCAAAATA  
121 AAAGAATGTA GTATATAGCA ATTGCTTTTC TGAGTTAT AAGTGTGTAT ATTTTAATT  
181 ATAACCTTTC TAATATATGA CCAAAATTG TTGATGTGCA GGTACGCGGG TGCTTACGAC  
241 CGTCAGTCGC GCGAGCGCGA GAATAGCAGT GCAGCAAGCC CAGCGACACA GCGTAGCGCC  
301 AACGAAGACA AGGCAGGCCA CCTTCAGCGC GAAGTCGAGC GCGACGGGGG CCGGTTCAAG  
361 TTCGTCGGGC ATTCAGCGA AGCGCCGGC ACAGTCGGCGT TCGGGACGGC GGAGCGCCCG  
421 GAGTCGAAC GCATCCTGAA CGAATGCCGC GCCGGGGCGC TCAACATGAT CATTGTCTAT  
481 GACGTGTTCGC GCTTCTCGCG CCTGAAGGTC ATGGACGCCA TTCCGATTGT CTCGGAATTG  
541 CTCGCCCTGG GCGTGACGAT TGTTCGACT CAGGAAGGCG TCTTCCGGCA GGGAAACGTC  
601 ATGGACCTGA TTCACCTGAT TATGCGGCTC GACGCGTCGC ACAAAAGAAC TTCGCTGAAG  
661 TCGCGAAGA TTCTCGACAC GAAGAACCTT CAGCGCGAAT TGGGCGGGTA CGTCGGCGGG  
721 AAGGCCTT ACGGCTTCGA GCTTGTTCG GAGACGAAGG AGATCACGCG CAACGGCCGA  
781 ATGGTCAATG TCGTCATCAA CAAGTTAGCG CACTCGACCA CTCCCCTTAC CGGACCCCTC  
841 GAGTCGAGC CCGACGTAAT CCGGTGGTGG TGGCGTGAGA TCAAGACGCA CAAACACCTT  
901 CCCITCAAGC CGGGCAGTCAGC AGCCGCCATT CACCCGGGCA GCATCACGGG GCTTTGTAAG  
961 CGCATGGACG CTGACGCCGT GCCGACCCGG GGCAGACGA TTGGGAAGAA GACCGCTTCA  
1021 AGCGCCTGGG ACCCGGCAAC CGTTATGCGA ATCCTTCGGG ACCCGCGTAT TGCGGGCTTC  
1081 GCCGCTGAGG TGATCTACAA GAAGAAGCCG GACGGCACGC CGACCACGAA GATTGAGGGT  
1141 TACCGCATTG AGCGCGACCC GATCACGCTC CGGCCGGTCG AGCTTGATTG CGGACCGATC  
1201 ATCGAGCCCG CTGAGTGGTA TGAGCTTCAG GCGTGGTTGG ACCGGCAGGGG GCGCGGCAAG  
1261 GGGCTTCCC GGGGGCAAGC CATTCTGTCC GCCATGGACA AGCTGTACTG CGAGTGTTGGC  
1321 GCCGTCACTGA CTTCGAAGCG CGGGGAAGAA TCGATCAAGG ACTCTTACCG CTGCCGTGCG  
1381 CGGAAGGTGG TCGACCCGTC CGCACCTGGG CAGCACGAAG GCACGTGCAA CGTCAGCATG  
1441 GCGGCACTCG ACAAGTTCTG TGCGGAACGC ATCTTCAACA AGATCAGGCA CGCCGAAGGC  
1501 GACGAAGAGA CGTTGGCGCT TCTGTGGAA GCGGCCGAC GCTTCGGCAA GCTCACTGAG  
1561 GCGCCTGAGA AGAGCGGCGA ACGGCGAAC CTTGTTGCGG AGCGCGCCGA CGCCCTGAAC  
1621 GCCCTTGAAG AGCTGTACGA AGACCGCGCG GCAGGAGCTT ACGACGGACC CGTGGCAGG  
1681 AAGCACTTCC GGAAGCAACA GGCAGCGCTG ACGCTCCGGC AGCAAGGGG GGAAGAGCGG  
1741 CTTGCCGAAC TTGAAGCCGC CGAAGCCCCG AAGTTGCCCC TTGACCAATG GTTCCCCGAA  
1801 GACGCCGACG CTGACCCGAC CGGCCCTAAG TCAGTGGTGGG GGCAGCGCTC AGTAGACGAC  
1861 AAGCGCGTGT TCGTCGGGCT CTTCGTAGAC AAGATCGTTG TCACGAAGTC GACTACGGGC  
1921 AGGGGGCAGG GAACGCCAT CGAGAAGCGC GCTTCGATCA CGTGGGCGAA GCCGCCGACC  
1981 GACGACGACG AAGACGACGC CCAGGACGGC ACGGAAGACG TAGCGGGCGTA G

## FIGURE 8

### Nucleotide sequence of pBPS EW051 T-DNA Region

#### Sequence Molecule Features:

Start	End	Name
3	217	Left T-DNA Border
225	259	attB
485	273	g7pA (terminator)
2288	519	<i>codA-aacCI</i> translational fusion gene
2898	2303	Nopaline Synthase Promoter
2925	3236	Octopine Synthase Promoter
3260	4267	tTA gene
4292	4558	Nopaline Synthase Terminator
4597	4933	Top10 promoter
4977	7007	$\phi$ C31 <sup>INT</sup> gene
7027	7221	Octopine Synthase Terminator
7253	8392	Super Promoter
8413	9405	<i>erGFP7<sup>INT</sup></i> gene
9411	9677	Nopaline Synthase Terminator
9690	9728	attP
9735	9880	Right T-DNA Border

#### Sequence:

1 TGGTGATTT GTGCCGAGCT GCCGGTCGGG GAGCTGTTGG CTGGCTGGTG GCAGGATATA  
61 TTGTGGTGTAAACAAATTGA CGCTTAGACA ACTTAATAAAC ACATTGCGGA CGTCTTAAT  
121 GTACTGAATT AACATCCGTT TGATACTTGT CTAAAATTGG CTGATTTCGA GTGCATCTAT  
181 GCATAAAAAC AATCTAATGA CAATTATTAC CAAGCAGGAT CACCGGTGCC AGGGCGTGCC  
241 CTTGGGCTCC CCGGGCGCGG CCCGGCAAT TCCCATCTTG AAAGAAATAT AGTTAAATA  
301 TTTATTGATA AAATAAGTCA GGTATTATAG TCCAAGCAAA AACATAATT ATTGATGCAA  
361 AGTTAAATT CAGAAATATT TCAATAACTG ATTATATCAG CTGGTACATT GCCGTAGATG  
421 AAAGACTGAG TGCGATATTAT GTGTAAATAC ATAAATTGAT GATATAGCTA GCTTAGCTCA  
481 TCGGGGGATC CTTAATCGAC TCTAGCTAGA ACGAATTGTT AGGTGGCGGT ACTTGGGTG  
541 ATATCAAAGT GCATCACTTC TTCCCGTATG CCCAACTTTG TATAGAGAGC CACTGCGGGA  
601 TCGTCACCGT AATCTGCTTG CACGTAGATC ACATAAGCAC CAAGCGCGTT GGCTCATG  
661 TTGAGGAGAT TGATGAGCGC GGTGGCAATG CCCTGCCTCC GGTGCTCGCC GGAGACTGCG  
721 AGATCATAGA TATAGATCTC ACTACGGGGC TGCTCAAACC TGGGCAGAAC GTAAGCCGCG  
781 AGAGCGCCAA CAACCGCTTC TTGGTCAAG GCAGCAAGCG CGATGAATGT CTTACTACGG  
841 AGCAAGTTCC CGAGGTAATC GGAGTCGGC TGATGTTGGG AGTAGGTGGC TACGTCTCCG  
901 AACTCACGAC CGAAAAGATC AAGAGCAGCC CGCATGGATT TGACTTGGTC AGGGCCGAGC  
961 CTACATGTGC GAATGATGCC CATCTCGAG AAACGTTGT AATCGATGCC TTCTGGCTGC  
1021 TCCAGATATA CGGTGGTTG TGCGGTTGT GTGCTGGCAA TCACCTTGCC GCCACGTACC  
1081 GAATAACGTA CGGAAACCTG ACGGCGCAGC GCATCAAACC CATTTCAGC CGGCAGGATA  
1141 ATCAGGTTGG CGCTGTTCTC GGCGGCAATG CCGTAATCCT GCAAATTCAA CGTCCTTGC  
1201 CTGTGGTGGG TGATTAAATT CAGGCCATCG TTAATCTGCC CGTAGCCCCAT CAACTGGCAA  
1261 ACATGCAGCC CCATATGCAG CACTTGCAGC ATATTGCCG TTCCCAGCGG ATACCACGGA  
1321 TCGAAGACAT CATCGTGACC AAAGCAGACG TTAATGCCGG ACTCCAGCAT CTCTTAAACG  
1381 CGCGTGTGTC CGCGACGTTT TGGATACGTA TCGAAACGTC CTTGCAGATG AATATTGACC  
1441 AGCGGGTTGG CGACAAAGTT AATACCGGAC ATTTTCAGCA AGCGGAACAG GCGTGAGGTA  
1501 TACGCCCCGT TATAGGAGTG CATTGCCGTG GTGTGGCTGG CGGTGACTCG CGCGCCCATG  
1561 CCTTCATGGT GCGCCAGGGC AGCAACGGTT TCGACAAAGC GCGACTGCTC GTCATCGATC  
1621 TCATCACAGT GAACGTCGAT GAGACGGTCG TATTTTGCG CCAGGGCGAA GGTTTTATGC  
1681 AGCGACTCCA CGCCGTATTC ACGGGTAAT TCAAATGCG GAATCGCCCC CACTACATCT  
1741 GCCCTAAGC GTAACGCCTC TTCCAGCAAC GCTTCACCGT TGGGATACGA CAAAATCCCT  
1801 TCCTGAGGGGA AGGGACGAT TTGCAGATCA ATCCACGGCG CGACTTCCTG CTTCACTTCC  
1861 AGCATTGCTT TCAGCGCAGT TAGCGTTGCA TCCGAAACAT CGACATGGGT ACGCACATGC  
1921 TGAATGCCGT TGGCAATCTG CCATITCAGC GTTIGCCATG CGCGTTGTTT CACATCGTCA

### FIGURE 8 CONTINUED

1981 TGGGTTAATA ACGCTTGCG CTCGGCCAG CGTCAATGC CTTCAAACAG CGTGCCGGAC  
2041 TGATTCCAGT TCGGTTGTCC GGCGGTTGC GTGGTGTCCA GGTGAATATG TGGCTCCACA  
2101 AACGGCGGTA TAACTAAACC TTGTCGGCA TCCAGGCTGT TTTCAGTTAT GGGCATCACG  
2161 CCGGATTGCG CATCAATGGC GCTGATTTC CCGTCCTGCA GATGAATCTG CCACAGCCCC  
2221 TCTTCGCCTG GTAACCGGGC GTTAATAATT GTTGTAAG CGTTATTGCA CACTGTTAGC  
2281 CTCCCCATGG AGATCTGGAT TGAGAGTGAA TATGAGACTC TAATTGGATA CCGAGGGGAA  
2341 TTTATGGAAG TCAGTGGAGC ATTTTGACA AGAAATATT GCTAGCTGAT AGTACCTTA  
2401 GGCGACTTT GAACCGCAGA TAATGGTTG TGACGTATGT GCTTAGCTCA TTAAACTCCA  
2461 GAAACCCCGG GCTGAGTGGC TCCTCAACG TTGCGGTTCT GTCAGTTCCA AACGTAAC  
2521 GGCTTGTCCC GCGTCATCGG CGGGGGTCAT AACGTGACTC CCTTAATTCT CCGCTCATGA  
2581 TCTTGATCCC CTGGGCCATC AGATCCTGG CGGCAAGAAA GCATCCAGT TTACTTTGCA  
2641 GGGCTTCCA ACCTTACCAAG AGGGCGCCCC AGCTGGCAAT TCCGGTTCGC TTGCTGTCCA  
2701 TAAAACCGCC CAGTCTAGCT ATCGCCATGT AAGCCCACG CAAGCTACCT GCTTCTCTT  
2761 TCGGCTTGC GTTTCCCTTG TCCAGATAGC CCAGTAGCTG ACATTGATCC GGGTCAGCA  
2821 CCGTTCTGC GGACTGGCTT TCTACGTGTT CCGCTTCCCT TAGCAGCCCT TGCGCCCTGA  
2881 GTGCTTGCAG CAGCGTGAAG CTTGGCGCGC CAAGCTGCA TGCCCGCTCT TAGCCGTACA  
2941 ATATTACTCA CCGGTGCGAT GCCCCCCATC GTAGGTGAAG GTGGAAATTAA ATGATCCATC  
3001 TTGAGACCAC AGGCCACAA CAGCTACCAAG TTTCTCAAG GTGCCACCAA AAACGTAAGC  
3061 GCTTACGTAC ATGGTCGATA AGAAAAGGCA ATTTGTAGAT GTTAACATCC AACGTCGCTT  
3121 TCAGGGATCC TTTTACCGA CAACTCATCC ACATTGATGG TAGGCAGAAA GTTAAAGGAT  
3181 TATCGCAAGT CAATACTTGC CCATTGATTG ATCTATTAA AGGTGTGGCC TCAAGGAGAT  
3241 CCCCGGGCCG GCAATTATA TGTCTAGATT AGATAAAAGT AAAGTGATTA ACAGCGCATT  
3301 AGAGCTGCTT AATGAGGTCG GAATCGAAGG TTTAACAAACC CGTAAACTCG CCCAGAAGCT  
3361 AGGTGTAGAG CAGCCTACAT TGTATTGGCA TGTAAAAAAAT AAGCGGGCTT TGCTCGACGC  
3421 CTTAGCCATT GAGATGTTAG ATAGGCACCA TACTCACTT TGCCCTTAG AAGGGGAAAG  
3481 CTGGCAAGAT TTTTACGTA ATAACGCTAA AAGTTTACA TGTGCTTAC TAAGTCATCG  
3541 CGATGGAGCA AAAGTACATT TAGGTACACG GCCTACAGAA AAACAGTATG AAACCTCTGA  
3601 AAATCAATTAA GCCTTTTAT GCCAACAAAGG TTTTCACCA GAGAATGCAT TATATGCACT  
3661 CAGCGCTGTG GGGCATTAA CTTAGGTGTT CGTATTGGAA GATCAAGAGC ATCAAGTCGC  
3721 TAAAGAAGAA AGGGAAACAC CTACTACTGA TAGTATGCCG CCATTATTAC GACAAGCTAT  
3781 CGAATTATTGATCACCAAG GTGCAGAGCC AGCCTCTTA TTCGGCCTTG AATTGATCAT  
3841 ATGCGGATTA GAAAAACAAC TTAAATGTGA AAGTGGGTCC GCGTACAGCC GCGCGCGTAC  
3901 GAAAACAAT TACGGGTCTA CCATCGAGGG CCTGCTCGAT CTCCCGGACG ACGACGCC  
3961 CGAAGAGGCG GGGCTGGCGG CTCCCGCCT GTCTTCTC CCCGGGGAC ACACGCGCAG  
4021 ACTGTCGACG GCCCCCCCGA CCGATGTCAG CCTGGGGAC GAGCTCCACT TAGACGGCGA  
4081 GGACGTGGCG ATGGCGCATG CCGACGCGCT AGACGATTTC GATCTGGACA TGTTGGGGGA  
4141 CGGGGATTCC CGGGTCCGG GATTACCCCC CCACGACTCC GCCCCCTACG GCGCTCTGGA  
4201 TATGGCCGAC TTGAGTTG AGCAGATGTT TACCGATGCC CTTGGAATTG ACGAGTACGG  
4261 TGGGTAGGGG GCGCGAGGAT CTCGAGCAGC TCGAATTTC CCGATCGTTC AAACATTGG  
4321 CAATAAAAGTT TCTTAAGATT GAATCCTGTT GCGGGTCTTG CGATGATTAT CATATAATT  
4381 CTGTTGAATT ACCTTAAGCA TGTATAATT AACATGTAAT GCGATGACGTT ATTATGAGA  
4441 TGGGTTTITA TGATTAGAGT CCGCAATTAA TACATTAAACCGATAGA AAACAAAATA  
4501 TAGCGCGAA ACTAGGATAA ATTATCGCGC GCGGTGTCACTATGTTACT AGATCGGGAA  
4561 TTCTTAATT AAGAATTGCA GCTCGGTACC GAGCTCGACT TTCACTTTCT TCTATCACTG  
4621 ATAGGGAGTG GTAAACTCGA CTTCTATTCTCTACT GATAGGGAGT GTAAACTCG  
4681 ACTTTCACTT TTCTCTATCA CTGATAGGGA GTGGTAAACTCGACTTTCACT TTTCTCTAT  
4741 CACGGATAGG GAGTGGTAACTCGACTTTCACT GATAGGGAGT GTAAACTCGAC TTTCACTTT  
4801 AACTCGACTT TCACCTTCTCTACTGATA TAGGGAGTGG TAAACTCGAC TTTCACTTT  
4861 CTCTATCACT GATAGGGAGT GTAAACTCG AGATAGAGTG ATCTAGTCTT CGCAAGACCC  
4921 TTTACGTATA TAAGGCCTTT CTAGACATT GCTCGAGCCC GGGGATCCAT ATGGCCATGG  
4981 CACAAGGGGT TGTGACCGGG GTGGATACGT AAGTTCTGC TTCTACCTT GATATATATA  
5041 TAATAATTAT CATTAAATTAG TAGTAATATA ATATTCAAA TATTITTTTC AAAATAAAAG  
5101 AATGTAGTAT ATAGCAATTG CTTTCTGTA GTTATAAGT GTGTATATT TAATTATAA  
5161 CTTTCTAAT ATATGACCAA AATTGTTGA TGTGCAGGTA CGCGGGTGCT TACGACCGTC

### FIGURE 8 CONTINUED

5221 AGTCGCGCGA GCGCGAGAAT TCGAGCGCAG CAAGCCAGC GACACAGCGT AGCGCCAACG  
5281 AAGACAAGGC GGCGCACCTT CAGCGCGAAG TCGAGCGCGA CGGGGGCCGG TTCAGGTTCG  
5341 TCGGGCATT CAGCGAAGCG CCGGGCACGT CGCGCTCGG GACGGCGGAG CGCCCGGAGT  
5401 TCGAACGCAT CCTGAACGAA TGCGCGCCG GGCGCTCAA CATGATCATT GTCTATGACG  
5461 TGTCGCGCTT CTCGCGCTG AAGGTATGG ACGCGATTCC GATTGTCTCG GAATTGCTCG  
5521 CCCTGGGCGT GACGATTGTT TCCACTCAGG AAGGCCTT CCGGCAGGGA AACGTATGG  
5581 ACCTGATTCA CCTGATTATG CGGCTCGACG CGTCGCACAA AGAACTCTCG CTGAAGTCGG  
5641 CGAAGATTCT CGACACGAAG AACCTTCAGC GCGAATTGGG CGGGTACGTC GGCGGGAAAGG  
5701 CGCCTTACGG CTTCGAGCTT GTTTCGGAGA CGAAGGAGAT CACGCGCAAC GGCGAATGG  
5761 TCAATGTCTG CATCAACAAG CTTGCGCACT CGACCACTCC CCTTACCGGA CCCTTCGAGT  
5821 TCGAGCCCCA CGTAAATCCGG TGGTGGTGGC GTGAGATCAA GACGCACAAA CACCTTCCTT  
5881 TCAAGCCCCG CAGTCAGCC GCCATTCAACC CGGGCAGCAT CACGGGGCTT TGTAAGCGCA  
5941 TGGACGCTGA CGCCGTGCCG ACCCGGGCG AGACGATTGG GAAGAAGACC GCTTCAGCG  
6001 CCTGGGACCC GGCAACCGTT ATGCGAATCC TTGGGGACCC GCGTATTGCG GGCTTCGCCG  
6061 CTGAGGTGAT CTACAAGAAG AAGCCGGACG GCACGCCGAC CACGAAGATT GAGGGTTACC  
6121 GCATTCAAGCG CGACCCGATC ACGCTCCGGC CGTCGAGCT TGATTGCGA CCGATCATCG  
6181 AGCCCCTGA GTGGTATGAG CTTCAGCGT GGTTGGACGG CAGGGGGCGC GGCAAGGGC  
6241 TTCCCCGGGG GCAAGCCATT CTGTCGCCA TGGACAAGCT GTACTGCGAG TGTGGCGCCG  
6301 TCATGACTTC GAAGCGCGG GAAGAATCGA TCAAGGACTC TTACCGCTGC CGTCGCCGGA  
6361 AGGTGGTCGA CCCGTCGCA CCTGGCAGC ACGAAGGCAC GTGCAACGTC AGCATGGCG  
6421 CACTCGACAA GTTCGTTGCCG GAACGCATCT TCAACAAAGAT CAGGCACGCC GAAGGGCAGC  
6481 AAGAGACGTT GGCGCTCTG TGGAAGCCG CCCGACGCTT CGGCAAGCTC ACTGAGGC  
6541 CTGAGAAGAG CGCGAACCGG GCGAACCTTG TTGCGGAGCG CGCCGACGCC CTGAACGCC  
6601 TTGAAGAGCT GTACGAAGAC CGCGCGGAG GCGCTACGA CGGACCCGTT GGCAGGAAGC  
6661 ACTTCCGAA GCAACAGGCA GCGCTGACGC TCCGGCAGCA AGGGGGCGAA GAGCGGCTTG  
6721 CCGAACTTGA AGCCGCCGAA GCCCCGAAGC TTCCCCTGA CCAATGGTTC CCCGAAGACG  
6781 CCGACGCTGA CCCGACCGGC CCTAAGTCGT GGTGGGGCG CGCGTCAGTA GACGACAAGC  
6841 GCGTGTTCGT CGGGCTCTTC GTAGACAAGA TCGTTGTAC GAAAGTCGACT ACGGGCAGGG  
6901 GGCAGGGAAC GCCCATCGAG AAGCGCGCTT CGATCACGTG GCGGAAGCCG CCGACCGACG  
6961 ACGACGAAGA CGACGCCAG GACGGCACGG AAGACGTAGC GCGTAGCTG CAGCTCGACG  
7021 CATGCCCTGC TTAAATGAGA TATGCGAGAC GCCTATGATC GCATGATATT TGCTTCAT  
7081 TCTGTTGTGC ACGTTGAAA AAACCTGAGC ATGTGTAGCT CAGATCCTTA CGGCCGGTT  
7141 CGGTTCATTC TAATGAATAT ATCACCCGTT ACTATCGTAT TTATATGAAT AATATTCTCC  
7201 GTTCAATTAA CTGATTGTCC AAGCTCCTG CAGGAAGCTT TGGCGGATC CTCTAGATT  
7261 GACGGTATCG ATAAGCTCGC GGATCCCTGA AAGCGACGTT GGATGTTAAC ATCTACAAAT  
7321 TGCCTTTCT TATCGACCAT GTACGTAAGC GCTTACGTT TTGGTGGACCC CTTGAGGAAA  
7381 CTGGTAGCTG TTGTGGGCCT GTGGTCTCAA GATGGATCAT TAATTTCCAC CTTCACCTAC  
7441 GATGGGGGGC ATCGCACCGG TGAGTAATAT TGTACGGCTA AGAGCGAATT TGGCCTGTAG  
7501 GATCCCTGAA AGCGACGTT GATGTTAAC TCTACAAATT GCCTTTCTT ATCGACCATG  
7561 TACGTAAGCG CTTACGTTT TTGGTGGACCC TTGAGGAAAC TGGTAGCTGT TGTGGGCCTG  
7621 TGGTCTCAAG ATGGATCATT AATTTCCACC TTCACCTACG ATGGGGGGCA TCGCACCGGT  
7681 GAGTAATATT GTACGGCTAA GAGCGAATTG GGCCTGAGG ATCCCTGAAA GCGACGTTGG  
7741 ATGTTAACAT CTACAAATTG CCTTTCTTA TCGACCATGT ACGTAAGCGC TTACGTTTT  
7801 GGTGGACCT TGAGGAAACT GGTAGCTGTT GTGGGCCTGT GGCTCAAGA TGGATCATT  
7861 ATTTCACCT TCACCTACGA TGGGGGGCAT CGCACCGGTG AGTAATATTG TACGGCTAAG  
7921 AGCGAATTG GCCTGTAGGA TCCGCGAGCT GGTCAATCCC ATTGCTTTG AAGCAGCTCA  
7981 ACATTGATCT CTTCTCGAT CGAGGGAGAT TTTCAAATC AGTGCAGCAAG ACGTGCAGTA  
8041 AGTATCCGAG TCAGTTTTA TTTTCTACT AATTGGTCG TTTATTCGG CGTGTAGGAC  
8101 ATGGCAACCG GGCTGAATT TCGCGGGTAT TCTGTTCTA TTCAAACCTT TTCTTGATCC  
8161 GCAGCCATTA ACGACTTTG AATAGATAAG CTGACACGCC AAGCCTCGCT AGTCAAAAGT  
8221 GTACCAAACA ACGCTTACA GCAAGAACGG AATGCGCGTG ACGCTCGCGG TGACGCCATT  
8281 TCGCCTTTTC AGAAATGGAT AAATAGCCCTT GCTTCTATT ATATCTCCC AAATTACCAA  
8341 TACATTACAC TAGCATCTGA ATTTCATAAC CAATCTCGAT ACACCAAATC GAAGATCCAA  
8401 GGAGATATAA CAATGAAGAC TAATCTTTT CTCTTCTCA TCTTTCACT TCTCCTATCA

### FIGURE 8 CONTINUED

8461 TTATCCTCGG CCGAATTGTA CGTAAGTTTC TGCTTCTACC TTTGATATAT ATATAATAAT  
8521 TATCATTAAAT TAGTAGTAAT ATAATATTTC AAATATTTT TTCAAAATAA AAGAATGTAG  
8581 TATATAGCAA TTGCTTTCT GTAGTTATA AGTGTGTATA TTTAATTAA TAACCTTTCT  
8641 AATATATGAC CAAAATTGT TGATGTGCAG GTACAATTCA GTAAAGGAGA AGAACCTTTTC  
8701 ACTGGAGTTG TCCCAATTCT TGTTGAATTAA GATGGTGATG TTAATGGGCA CAAATTTC  
8761 GTCAGTGGAG AGGGTGAAGG TGATGCAACA TACGGAAAAC TTACCCCTAA ATTATTTGC  
8821 ACTACTGGAA AACTACCTGT TCCATGGCCA ACACCTGTCA CTACTTCAC TTATGGTGT  
8881 CAATGCTTT CAAGATAACCC AGATCATATG AAGCGGCACG ACTCTTCAA GAGGCCATG  
8941 CCTGAGGGAT ACGTGCAGGA GAGGACCATC TCTTCAAGG ACGACGGGAA CTACAAGACA  
9001 CGTGTGAAG TCAAGTTGA GGGAGACACC CTCGTCAACA GGATCGAGCT TAAGGGAATC  
9061 GATTCAAGG AGGACGGAAA CATCCTCGGC CACAAGTTGG ATACAACTA CAACTCCCAC  
9121 AACGTATAAC TCA CGGCAGA CAAACAAAAG AATGGAATCA AAGCTAACTT CAAAATTAGA  
9181 CACAACATTG AAGATGGAAG CGTTCAACTA GCAGACCATC ATCAACAAAA TACTCCAATT  
9241 GCGGATGGCC CTGCTTTT ACCAGACAAC CATTACCTGT CCACACAAATC TGCCCTTCG  
9301 AAAGATCCC ACGAAAAGAG AGACCACATG GTCTTCTTG AGTTGTAAC AGCTGCTGGG  
9361 ATTACACATG GCATGGATGA ACTATACAAA CATGATGAGC TTAAAGAGCT CGAATTCCC  
9421 CGATCGTTCA AACATTGGC AATAAAGTTT CTIAAGATTG AATCCTGTT CGGGTCTTGC  
9481 GATGATTATC ATATAATTTC TGTTGAATTAA CGTTAAGCAT GTAATAATTAA ACATGTAATG  
9541 CATGACGTAA TTTATGAGAT GGGTTTTAT GATTAGAGTC CCGCAATTAT ACATTAAATA  
9601 CGCGATAGAA AACAAAATAT AGCGCGCAA CTAGGATAAA TTATCGCGCG CGGTGTCATC  
9661 TATGTTACTA GATCGGGAAAT TCGCGATCGC CCCAACTGGG GTAACCTTG AGTTCTCTCA  
9721 GTTGGGGGAG ATCTGATTGT CGTTCCCGC CTTCAGTTA AACTATCAGT GTTGACAGG  
9781 ATATATTGGC GGGTAAACCT AAGAGAAAAG AGCGTTATT AGAATAATCG GATATTAAA  
9841 AGGGCGTGAA AAGGTTATC CGTCGTCCA TTTGTATGTC

## FIGURE 9

### Nucleotide sequence of *Arabidopsis thaliana* G4H promoter region

1 TGTAATGAT AGGGATTGAA ACATCATCCT ATCGTTGACC AAAAATTCA CTGCGTGCTA  
61 TATAAAATAC TATATATGTT ACCCTTAAAC TGATGAAAAT GTAAAGAGAC AAGGCAGCAC  
121 CGTTTATCAT CAGACCAGTT TCGAGAGTGT TCCTGCATCG TTGGGCTCCC TCCTCAATT  
181 TGTCTACGTG ATTATATATC ATATCGTCTA CAAACAAAAT AAATACAATT CTATCATATG  
241 AATATGTGAT CATCGATGAT CGATCAATAT ATGTTTCGA GGTGACGTAT ATAGTATATT  
301 TCCCTAGAGA CGGCGAAGAA CATGATATCT CTGCATGCC CCAATCAAAT CTTTACACTT  
361 CATCCTCTT CGTTACTTGT TCAGTTGTC CTTCTAAC CCGACAACCC TTAATTGTA  
421 TTTCTATATT AGATCGAAAT ATCTCATTTG TGATAAATAA AATAAAAAAA ATCAAAGAAA  
481 GCTATAGAGA AGCTGCGTGC ATGCATGGGT TGGCGATGTT TGGCTTGTGA TGTITGGCTT  
541 GTTATGTGGC ATTATCTGTA TGTATATTAC CCTAAATCAC ATCTACGACA TTICCTCGA  
601 TCTTCAAAAT ATGCCAGCAA TCTTCATGTT TCCTCATATC TCTTAACATT GGAAAATGTC  
661 TTTTGACCTC TTTTGATGTA TTTTAAATTA CTTCGAGCTC ATCTATATTAA CAAATCATTC  
721 ATGGTGAATT ATTGTCAGC CAATAGAATA GAAATCTGAA TATAATGTGT ACCACATCTT  
781 TTATGTAATT TATACGATAT TCTTTTCCT GAGAATGATC AAATAACAAC ATGCATGAAT  
841 TGCTGCCAGA AAACGTCAGA TTGATCAGTT ATCACTACAA TTATCAATTAA ACTAGTAAAT  
901 AGTATCAAAA TGTACGTAGT GCCCATCTAT AGCTAGCTAA GGAGGACTCC GGATGTAGAG  
961 AAAAGCTAAA ATGTGACTTG CTAGAGTTGT ATTATATTGA ATTTCCTAA CTAATAGTAT  
1021 CTTTTTACA GATAATAATT TCCGGAAAAC CTATTAGATG TATAGATATA ACAATAAGCA  
1081 TCGATACCAA CCTTTTACTT CCAAAAAAAA ATAAAAAAA AATGCCAAGA TGAGATAATT  
1141 TTGTCAATT CAATTAGTGG GAAAATAACA ATTGTCGTGT TATTTCGAA CCAACGCATC  
1201 TCAGTGAATG ATTCCCAGT TCTTAAGATT TTAGGACATA CTTTCCAGT AACATCTAAT  
1261 CCGTTGGGC ATAAACAAGA CAATTGTAG TTATGTACAT TTCTTAGTGA TGTGTGTTGA  
1321 AAAGATATGA ATCAATGAGG TCCGACATAT TTTGTCAATA CGTTAGTGGT GTTTCAAAAT  
1381 AAATTTITAG TATATATATT AAAATAAGAC CAAAGGATAG GCTTTAGTGG TGTTTCAGGT  
1441 ATAGTTTAA TAATCAATT CAAAATAAGTC GAAAGGATAT GTAAGATAAG CGTTATTTC  
1501 ACGTGGATCA TTATCAACCA TGTAAAAAC GCATTCAAC TCCTAGATGT GTTGTAGTT  
1561 ATATATGTC CAAATGGAAT CGACCCAACA GAAAAGAGA AAAAAACGTA AAAGGTTATG  
1621 CGATTCCAGG GACGTCTCAT ATATATATAT ATTCCGATGAA AATATAAATA TAATTATCGT  
1681 GGTCTGTGAC AATAAAATATG GAAATAGATG TGGAAATCAT GATCATGTGA AGAAGAAGAA  
1741 GAACACGTGC AGATGAAC TGCAAAATGATAA TAATGTGAT GTCATGAGT TATGTAATT  
1801 TGTGTATTAT CTACGTGTT TCCATACATA CATATATAAA TCTTATATTAA CTITATGGTT  
1861 TTGTGTAAGGAGTACGTAG CATCAATAAT TGTGATTGTT TGCCATAAAC AGACAAC  
1921 TTGTAAACGGT ATAAGGCTTG GCTCTCATGA TAAAATGATA ACCCTTTTT TCGTCGGAGA  
1981 CAGACAAACG CATAAAATCAC TAATTCTAAA CCGAGATGAT TGTGATTGTT TTTGCCATAT  
2041 GCATAACTAG AATCTTCAGT TAATATTAAT TTTGGTGTGTT TTCGATCGAA TAAAAAAA  
2101 TAAACATTGC AATATTCGA AATTGTCGT CTTCTTTTATAACACTAG CAAGTGAGAG  
2161 GCTGAGAGCC AAGTGGAACG TTAAAAGACA ACATTAGATA TATATTATAT ATTGCTAAAT  
2221 CTGTATTATT TCTTTTAAC ATACGCAACT TTGATTGGA AATCGTAAGT CGAAGGAAGG  
2281 GCCTCGATTGTT ATGACGTACG CTTCTGTGCCA AACAATTCT CTITAGTTGA GGCCGGGGAA  
2341 GACGAGTTG TTGTAGTGA GCGATGCCAT GGCATCAATG AACCTCCAAA GGCCATATGT  
2401 TCTGTTAAAG GCTATTITAG TTGTTAATTT TGTCTGATT AACTCAACCA CATGTTAAAT  
2461 CAGATATCAT GTTAAACGAT ATTGTTTTT AAAACAAAATG ATTATCATAA AACGAAATIT  
2521 ATGATGAAAC ATATATAATC TTATCTTGT TTAAGTATGT AATCTTGTGA TGTTGTATA  
2581 CGCCTGCAA ATCAAAAAAC TAGTTGCTGT TTGTTGGCATT GTGTTACGA AATATTATT  
2641 AATATTITAA ATTAATTAAA TAAATGTTCT TATTCTCAA CAGGAAACAA TATGTATT  
2701 CTTCTTTAT AAAATTACAA TGAATTATTG TTGTTAAGCT GTCTATTCC AAGAAACAAA  
2761 ACACAAAAAT GATAAAATTAA TAATAGTCAC ATAACCTGTC TTACAAAAA AAAAGAAAA  
2821 GCGAAAAGAA ATGTGACAAAC AGAAAATGGT TTGATAACC AATAAGAATC GACAAAAAA  
2881 AAACCTTACTC CACATATACT CTTCTCTTCA CTCTTCAGTC TTCACTATTG AGTCTCGAGT  
2941 ATTTCACCGA TCTATAAATA CACTCTTCTT CTCCACAAA AGTATCATAT CATAACAAAA  
3001 ACATAAAGCC AAAATATAAA CACATAAGCC TTTTA